

DUEL
(‘Disfluencies, Exclamations and
Laughter in dialogue’ project)
Transcription and Annotation Manual

University of Bielefeld, Germany
Université Paris-Diderot (Paris 7), France

V1.0, April 27, 2015

Contents

1	Preface	3
2	Work-flow and technical set-up	4
2.1	Basics	4
2.1.1	Detailed listening: Extracting separate sound channels and slowing down sections of the audio in Praat	5
3	Task specific conventions	6
3.1	Transcription Tiers / Segmentation Levels	6
4	Segmentation Guidelines: general	6
4.1	Utterances	7
4.2	Turns	7
5	Transcription Guidelines: general	8
5.1	Disfluencies	9
6	Segmentation Guidelines: German	11
7	Transcription Guidelines: German	11
7.1	Utterance tiers	11
7.1.1	Orthography	11
7.1.2	Capitalization	11
7.1.3	Edit terms and filled pauses	12
7.1.4	Repair disfluencies	12
7.1.5	Non-verbal utterances	12
7.1.6	Non-standard German	13
7.2	English translation tiers	13
8	Segmentation Guidelines: French	14
9	Transcription Guidelines: French	14
9.1	Orthography	14
9.1.1	Filled pauses	14
10	Segmentation Guidelines: Chinese	15
11	Transcription Guidelines: Chinese	15
11.1	Utterance tiers	15
11.1.1	Orthography	15
11.1.2	Filled pauses and editing terms	15
11.2	Disfluencies	16
11.3	Pronunciation variations and words in languages other than Chinese	16
11.4	English translation tiers	16

1 Preface

This document is the manual for the transcription of dialogue data in the DUEL (“Disfluencies, Exclamations and Laughter in dialogue”) project between the University of Bielefeld and Université Paris Diderot (Paris 7).

This is a modification and extension of David Schlangen and Raquel Fernández’s Potsdam Dialogue Corpora transcription and annotation manual (Schlangen and Fernández, 2008).

2 Work-flow and technical set-up

2.1 Basics

We use **Praat** (Boersma, 2002) for the transcription work. **Praat** is a program for phonetic analysis but also works well for transcribing speech. **Praat** is free and there are Versions for Mac, Windows, Linux, and various UNIXes.

The audio file to be transcribed needs to be in WAV or AIFF format. You should have been provided with files in one of these formats. If however your file is in mp3 format, on Mac you can use **iTunes** to convert it to AIFF; on Windows there are a lot of free converters, for example **Mp3Decode**. On Linux Systems you can use **mpg123** for this step.

To start transcribing, you need to load the audio file into **Praat**. It can be useful to load the audio file in two different ways: both channels as one mono file (**Read→Open long sound file...**) and the two channels separately (**Read→Read two sounds from stereo file...**). For another way of extracting multiple channels see Section 2.1.1 below. This way you can choose to listen to both channels together or to either of them separately, depending on what you find easier for the transcription at each point. For instance, when the speakers overlap, you may find it easier to listen to them separately to transcribe the utterances; while to decide on turn boundaries you need to play both channels at the same time.

The transcription will be stored in a separate file, henceforth this file is called the **TextGrid** file. You will either have been provided with pre-configured **TextGrid** files or will have to create your own; see next section. Don't forget to save frequently while you are working.

The tasks for transcription that should be carried out are a) to *segment* the contributions of each speaker into utterances (see Section 4.1 below), b) to *transcribe* these utterance (Section 5), and c) to *segment* the contributions (or, equivalently, to group the utterances) into turns (Section 4.2).

If you have been provided with a **TextGrid** for your transcription task, load it. (Load the audio file as described above, load the **TextGrid** with **Read→Read from file...**, select both the audio object and the **TextGrid** object and click **View & Edit**.)

If not, create a **TextGrid** object and tiers as follows. Select the audio object, click **Annotate→To TextGrid**, enter the tier names according to the task at hand being described in section 3 and a tier labeled 'comments'.

If not transcribing an interaction in English, you will need to create separate tiers for an English paraphrase translation for each participant, which align with the utterance segments. Details of the translation tiers will be given in the language specific guidelines.

2.1.1 Detailed listening: Extracting separate sound channels and slowing down sections of the audio in Praat

If you would like to extract separate channels to focus on a given channel for transcription and annotation, do the following:

1. Go to the Praat objects window, select the long sound file which is the audio file you are transcribing.
2. Press the **Query** button, and select **Query time domain**→**Get total duration**. You will get a window with the length of the file in seconds, copy the number (without the 'seconds'). Close this window.
3. In the Praat objects window, go to **Extract part...**, then paste the number you have just copied into the right-hand window (the left one should be 0.0 as it indicates the start time). Press **OK**. Wait a few seconds for it to extract the part.
4. In the Praat objects window, select the final sound that has just been created, and click on **Convert-**, then select **Extract all channels**. Wait a few seconds for this to complete.
5. The channels from the file will have been created and are at the bottom of the objects list.
6. Select each extracted channel and from the **Save** menu, click on **Save as WAV file**
7. Open one of the WAV files as a long sound to isolate that channel, along with the TextGrid you are using in the objects window. Then, as usual, click **View & Edit**.

If you would like to slow down a particular interval of audio to better transcribe or annotate it do the following:

1. Select the interval you want to analyse.
2. Go to **File**→ **Extract selected sound (preserve times**
3. Go to the Praat objects window
4. Select the sound you have just created called 'untitled' at the bottom of the objects list
5. Click on **Manipulate-**
6. Click on **To manipulation...**, leave all settings as they are and click **OK**
7. Go back to the Praat objects window and click on the object called 'Manipulation untitled'
8. Click **View & edit**
9. In the manipulation window click on the **Dur** menu, and select **Add duration point at cursor**
10. The red dot in the lowest window **Duration manip** indicates the duration (and therefore speed) of the clip, moving it up will increase duration (and slow the

clip down). Move this point up to the level you need.

11. To listen to this interval at the new speed click **Visible part** [number]

3 Task specific conventions

Here we advise on the task-specific conventions of tier labeling.

3.1 Transcription Tiers / Segmentation Levels

For every transcription you should enter the following tier names into the field **Tier names**: **A-turns** **A-utts** **B-turns** **B-utts** corresponding to the utterances and turns of the two dialogue participants, leave the field **Point tiers** empty. Now select both the audio object together with the new **TextGrid** object and click **Edit**.

Either way, you should now have open a **TextGrid** that contains a minimum of 5 tiers:

1. A-turns: tier containing the turn boundaries for participant A
2. A-utts: tier used for segmentation and transcription of A's utterances
3. B-turns: tier containing B turn boundaries
4. B-utts: tier used for segmentation and transcription of B's utterances
5. Comments: tier for noting down any queries about the transcription at a particular point

Tiers 2 and 4 are used to transcribe utterances. In these tiers transcribers need to mark utterance boundaries, and then for each utterance interval transcribe the utterance in question (see the guidelines below).

Tiers 1 and 3 are used to mark turn boundaries. Inside the turn intervals, just write an **A** when the interval corresponds to the A's turn in tier 1, and a **B** when it corresponds to B's turn in tier 3 (i.e., don't repeat the textual transcription here.) See the turn segmentation guidelines above and below.

Marking utterance and turn boundaries should be done as precisely as possible.

Each task has a separate convention for selecting which tier corresponds to which participant in the dialogue. In the Dream Apartment (DAP) setting, simply use the participant who speaks first as **A**, and the other as **B**, however for the interview task this should be done as follows:

Task	Interview	Interview
Role	Interviewer	Interviewee
Tier label	A	B

4 Segmentation Guidelines: general

Before beginning segmentation, please refer to the language specific segmentation guidelines below (e.g. Section 6 for German, etc.) as you will need to create ad-

ditional tiers before segmenting.

4.1 Utterances

Our notion of utterance basically follows Meteer et al. (1995)'s concept of the *slash unit*: "A slash-unit is maximally a sentence but can be a smaller unit. [...] Intuitively, slash-units below the sentence level correspond to those parts of the narrative which are not sentential but which the annotator interprets as complete." (See therein for more details.). However note that as will be explained in Section 5, for restart repair disfluencies, we mark these specially if there is no silence, rather than coding the utterances as two separate units.

Some general notes:

- The general principle to follow is: segment more rather than less.
- Cut off leading "okay", "right", "so"s, first proper sentence is definitely an utterance. If there were several "okay", "right"s etc., group them according to intonatory boundaries.
- Leading "so": if it can be replaced by "hence", don't segment it off. (And so on—if you can paraphrase particles, they belong to the sentence.)
- Also separate leading "e:m" (filled pauses).
- "e:m" is part of utterance if it comes within a coherent syntactic structure (modulo disfluencies).
- "e:m" at end of abandoned utterance is also part of that utterance.
- Interruption by parenthetical utterance: all parts are individual utterances.
- The only exception of "split off as much as possible" principle: restart disfluencies are part of their utterance, are marked up with the appropriate disfluency markup explained in Section 5.1.

4.2 Turns

All (more or less) continuous stretches of speech by one speaker until other speaker takes over (some pauses are allowed within the stretch) are one turn. Exception: if it really sounds as if the speaker finishes and waits for other DP, and only resumes speaking after a silence in which other DP does not start speaking, this should be annotated as two consecutive turns by the same speaker.

The following is an example of a single turn with four utterances (stars indicate turn boundaries, pipes utterance boundaries):

```
*|Yeah, that's right| |Now (I've got this + . the shapes I've  
got here )| |The bottom right hand side . I think it's the  
elephant's leg| |I've got this L shape that is an L shape  
facing the wrong way|*
```

The next example shows two consecutive turns by the same speaker:

*|The longest part is horizontal, the three blocks| |And then
 (it's got e:m + the first block of these is got) a block on top
 of that . and then a block to the left of that block|* *|So|*

The turn boundaries should exactly coincide with the utterance boundaries on their left and right hand extremes, so care should be taken in segmentation to do this.

5 Transcription Guidelines: general

Please conform to the following guidelines when transcribing utterances.

Tags. The following tags can be used where appropriate:

Tag	Description	Example
<self> </self>	self directed talk	<self>hmm how can I do this</self>
<laughter/>	laughter bout	that's cool <laughter/>
<laughter></laughter>	laughed speech	<laughter>yeah</laughter>
<laughterOffset/>	sharp breath after laughter	that's cool<laughter/><laughterOffset/>
<breathing/>	audible breath	that's cool<breathing/>
<unclear/>	unclear material, left out	I'm <unclear/> this
<unsure></unsure>	inconfident transcription	to <unsure>the side</unsure>
<english></english>	English word(s) (within German/French/Chinese dialogue), though not loan words that are frequently used	genau <english>yes</english>

Prosodic issues. The following symbols can be used where appropriate.

Symbol	Description	Example
:	lengthening	eh:m:
?	rising tone	do you mean this one ?

Silent pauses. For a typical speaker, please transcribe pauses of short, medium and long length according to this rule of thumb:

. (0-0.33 sec) / .. (0.33-0.66 sec) / ... (> 0.66 sec)

For example: *They go . on top .. of the other one*

Having said this, if speakers have slower or faster rates of speech than an average speaker, it will be up to the judgement of the transcriber as to whether it was a short, medium or long pause.

Non-standard variants and mis-speaking For words which are pronounced in a non standard way, use the convention `<v=" [normal form] ">[pronounced form]</v>` (e.g. `<v="camera">cameral</v>`). Be careful not to over-correct speech here– if this is a standard form in speech, then do not use this and transcribe using standard orthography. See the individual sections for each language for more details. For obvious mis-pronunciations use `<m=" [normal form] ">[pronounced form]</m>` (e.g. `<m="angry">angly</m>`).

Non-verbal utterances Typically, monosyllabic functional utterances we use `hm` (with no brackets) and for the disyllabic equivalent we will use `mhm`. This may vary for different languages– when it does this will be explained in the appropriate language-specific sections below. As regards other non-verbal vocalizations, apart from laughter and breathing and these functional utterances, for non-linguistic contributions such as coughing, sneezing and lip-smacking, use a `<nonverbal/>` tag.

5.1 Disfluencies

The following overviews the disfluency annotation conventions you should follow. Please take special care to mark all of these that you can hear as accurately as you can, and when needed use channel separation and slowing down the audio as in Section 2.1.1 if you are initially unsure about what annotation tags are best.

Editing terms and filled pauses While the inventory of edit terms will differ from language to language, their general definition are words which are not part of standardised syntax for the language but perform some editing or discourse function. These will be marked within curly braces, e.g. `I { you know } like her`. With editing terms we distinguish different types, following Meteer et al. (1995), presently only distinguishing filled pauses from the rest by including an ‘F’ e.g. `I {F uh } went fishing`.

Exclamations We mark up exclamations in a similar way, though use an ‘X’ to distinguish them from edit terms and filled pauses. e.g. `{X God } That’s a pain.`, `{X Wow }` etc.

Repair disfluencies. In terms of the structural level of repair disfluencies, we largely follow Meteer et al. (1995) in tagging the structure of utterance-medial repair disfluencies (repetitions, substitutions and deletes) according to this scheme:

$$(\text{LHS} + \{ \text{Editing Term} \} \text{RHS})$$

where the editing term is optional, and the RHS is not present in a deletion or canceling type repair- i.e. there is no candidate RHS that makes sense as a substitution for the

LHS. For instance while the first example clearly has a RHS component, the second one does not, and is a delete.

```
I went to ( the: + {F um } the ) garden
She said ( there + ) I could go tomorrow
```

Chaining and nested repair disfluencies When a speaker has trouble and is disfluent in a burst where the RHS of one repair forms the LHS of another, this is a *chaining* repair disfluency and should be given an embedded structure, where you should assume rightward-branching:

```
( I + ( I + I ) ) want to go to Berlin
```

When the disfluency is *nested* with a longer repair this can be annotated as follows:

```
(In any case + in ( any + any ) case) I want to go to Berlin
(In (any + any) case + in any case) I want to go to Berlin
```

Restart disfluencies and abandoned utterances We depart from the Switchboard disfluency conventions in (Meteer et al., 1995) in that we also mark *restarts* structurally within the same utterance, rather than separating them as different utterance units within the same turn. This follows the standard representation of delete repair disfluencies, as in the below where the restart begins from ‘yesterday’:

```
(I + {F uh }) yesterday someone said yes to that
```

A common decision to make when annotating a restart is whether you think it is restarting and repairing the turn (and hence has some similarity in function to the original attempted LHS), or it is an *abandoned utterance* followed by a completely new utterance. In cases with an interruption from the speaker of the abandoned utterance’s dialogue partner more clearly signals this, however the abandonment may be self-initiated, followed by a long pause, and then resumption with a new utterance. In these cases, mark the abandoned turn with a - and segment the next utterance anew:

```
|I guess that’s -| |Yesterday someone said yes to that|
```

The other case of abandoned utterances are when there is no long pause, but the utterance is abandoned for a completely different type of dialogue act, e.g. here, where an acknowledgement is in the new utterance.

```
|I guess that’s -|Yes, correct.|
```

Partial words and mid-word pauses For partial cut-off words use a ‘-’ after the last recognizable syllable has completed, flush to the partial word, with the best approximation for that syllable in the following bracketing convention, where ‘p’ stands for ‘partial form of’: e.g. I (<p= ‘tried">tr-</p> + tried) to like Mary . If you are unsure what the partial word is, simply tag with a <p> opening tag.

For mid-word pauses and fillers where the word is resumed use a ‘-’ prefixed on to the resumption part of the word, in these cases just mark the partiality of the words. e.g. It’s <p>fan-</p> {F uh } <p>-tastic</p> .

6 Segmentation Guidelines: German

Before segmenting utterances, in addition to the 5 tiers already created, you will need 2 more tiers, which consist of English paraphrase translation tiers for each participant, labeled as follows:

1. A-de_en: tier containing an English translation paraphrase of A’s utterances
2. B-de_en: (the same as above for A) utterances

You should now have a total of 7 labelled tiers: A-turns, B-turns, A-utts, B-utts, A-de and B-de and Comments.

7 Transcription Guidelines: German

7.1 Utterance tiers

The basic principle is to be ‘close, but not too close’ to the original when transcribing into the utterances tiers. You should leave in non-standard grammatical constructions and slang words without extra annotation, however, for non-standard pronunciation or lexical variants you should use the non-standard German annotation conventions as described below.

7.1.1 Orthography

Standard German spelling should be used for all forms where possible, except in the cases of partial words and filled pauses (see section 5), even if phonetically different from High German.

7.1.2 Capitalization

Capitalization for nouns and utterance-initial words should be consistent.

7.1.3 Edit terms and filled pauses

When transcribing filler sounds, please use the transcription that seems most adequate from the following list (remember these should always be transcribed with an {F } surrounding them):

{F äh} / {F ähm} / {F hm}

We also have other edit terms, these include the following (should be transcribed with a { } surrounding them:

{also}

7.1.4 Repair disfluencies

Please refer to Section 5 for details on repair disfluency annotation. Here we have a selection of some more complex disfluency examples with the audio attached. Firstly, the following are examples of a restart disfluency at the beginning of the first utterance. Note these should be part of the same utterance:

```
| (können ja auch zw- + {na}) wir dürfen ja nur eine Küche machen |
|({also} +) aber so in dem .. |
| (genau nicht verwinkelt so dass +) und . breit genug dass man noch was reinstellen
kann|
```

Here we have a nested repair at the beginning, which forms part of a chaining repair (notice the embedded structure of the bracketing. The second set of embedded repairs from “das Wohnzimmer . mit dem” again has a nested repair in the reparandum phase within a longer repair. Notice the filled pause and “also” are bracketed as edit terms. Also notice that while the first partial word disfluency is marked <p="können">kö-</p> this is not done for Wohn- und Ess-Zimmer as this is a fluent German construction.

```
|((das <p="können">kö-</p> + da können ) wir dann + .. (da können wir dann + .
da können wir dann )) besser vielleicht noch {F ähm} (das ( Wohnzimmer . mit
dem + ... Wohn- und Ess-Zimmer ) . + {also} das Wohnzimmer mit der )
Wohn-Küche .. trennen oder nur so'n Durchgang machen|
```

Finally, another restart with a filled pause and an interregnum phrase “ja also” before the repair phase on the RHS of the main bracketing and a delete utterance finally:

```
| ( (<p="Wohnzimmer">Wohn-</p> +) . {ja also} (die + (die + das)) {F äh} ...
Wohnzimmer . (dementsprechend dann irgendwie +))|
```

7.1.5 Non-verbal utterances

For monosyllabic functional utterances we use hm (with no brackets) and for the disyllabic equivalent we will use mhm. For other non-verbal vocalizations, apart from

laughter and breathing and these functional utterances, such as coughing, sneezing and lip-smacking, use a <nonverbal/> tag.

7.1.6 Non-standard German

While in the utterances tiers you should use standard German spelling wherever possible, phonetic forms may often differ significantly to their standard forms, particularly in the case of common contractions. The basic rule is: if it is conventional in conversation to have this form as a complete word, such as “ne” at the end of an utterance, then transcribe it as it is, as opposed to another formal written German form. For uses of English words in conversation, if they are conventionalized, i.e. ‘cool’, ‘Whirlpool’, do not annotate them as English words as described in Section 5, but only if there is a quotation in English or a definite switch to speaking English.

When encountering contractions and non-standard pronunciation transcribe with approximations to the speaker’s pronunciation such as those within the XML-style tags in the right-hand column in Table 1. These shortened forms should be annotated in the form of the right-hand column, where the v stands for ‘variant of’, and on the right hand side of the = sign you should write the standard German form. This is not an exhaustive list, though is a list of frequently occurring forms. The missing of the final schwa ‘e’ sound in forms such as *glaub*’ is common and other verbs not in this list should be annotated the same way. Also note that forms beneath the line in the table are not tagged in this way, as they are standardized tokens in spoken speech.

7.2 English translation tiers

For English translation, upon segmenting utterances and having transcribed the German appropriately, write a paraphrase to the best of your ability in English in the appropriate *de_en* tier. We ask you to use standard English spelling and to preserve the disfluency markings you have made in German, using the same bracketing scheme (Section 5) and replace or maintain filled pause orthography with the following English correspondences:

{F äh} / {F ähm} / {F hm} (for German utterances tiers)
{F uh} / {F um} / {F hm} (for English translation *de_en* tiers);

The non-verbal utterances as described in Section 7.1.5 will have the same orthography as for the German utterance tiers.

form heard	standard form	transcription convention
aufer	auf der	<v="auf der">aufer</v>
det	das	<v="das">det</v> brauchen wir nicht
das	das ist	<v="das ist">das'</v> ja blöd
s'	das ist	<v="das ist">s'</v> ja blöd
dohoch	doch	<v="doch">dohoch</v>
'n	ein	da steht <v="ein">'n</v> Auto
'ne	eine	wir bauen da <v="eine">'ne</v> Sauna rein
'nen	einen	er hat <v="einen">'nen</v> Garten
'ner	einer	<v="einer">'ner</v> Wohnküche
für's	für das	das ist <v="für das">für's</v> Wohnzimmer
ham	haben	wir <v="haben">ham</v> Glück
glaub'	glaube	ich <v="glaube">glaub'</v> das nicht
halt'	halte	ich <v="halte">halt'</v> mich im Zimmer auf
is'	ist	das <v="ist">is'</v> mir egal
lauf'	laufe	ich <v="laufe">lauf'</v> rüber
machen's	machen es	wir <v="machen es">machen's</v> morgen
mein'	meine	ich <v="meine">mein'</v>, dass das gut ist
nehm'	nehme	dann <v="nehme">nehm'</v> ich das
nich	nicht	das brauchen wir <v="nicht">nich</v>
isses	ist es	<v="ist es">isses</v>
haste	hast du	<v="hast du">haste</v>
is	ist	<v="ist">is</v>
'kay	okay	<v="okay">'kay</v>
weisste	weisst du	<v="weisst du">weisste</v> wo es ist?
wa	wir	das gönnen <v="wir">wa</v> uns
wär'	wäre	das <v="wäre">wär'</v> gut
würd'	würde	<v="würde">würd'</v> ich sagen
naja	naja	naja ... ich weiß nicht
ne	ne	das Wetter ist schön, ne?
nee	nee	nee das brauchen wir nicht
nix	nix	da ist nix

Table 1: Non-standard German forms and their transcription conventions

8 Segmentation Guidelines: French

9 Transcription Guidelines: French

9.1 Orthography

9.1.1 Filled pauses

10 Segmentation Guidelines: Chinese

Before segmenting utterances, in addition to the 5 tiers already created, you will need 2 more tiers, which consist of English paraphrase translation tiers for each participant, labeled as follows:

1. A-ch_en: tier containing an English translation paraphrase of As utterances
2. B-ch_en: (the same as above for A) utterances

You should now have a total of 7 labeled tiers: A-turns, B-turns, A-utts, B-utts, A-ch_en, B-ch_en and Comments. To make sure that the turn boundaries match the left and right boundaries of the utterances within the turn, select a turn interval and press ctrl + the tier number for the relevant utterance tier, to copy the turn boundary onto the utterance tier.

11 Transcription Guidelines: Chinese

11.1 Utterance tiers

The basic principle is to be close, but not too close to the original when transcribing into the utterances tiers. You should leave in non-standard grammatical constructions and slang words without extra annotation. For filled pauses, certain discourse markers and certain non-standard pronunciations, follow the instructions below.

11.1.1 Orthography

Use Chinese characters where possible, except in the cases of pause fillers and non-standard discourse particles. See below.

11.1.2 Filled pauses and editing terms

For filled pauses, use the form {F }, e.g {F en1}. Write down the Pinyin of the words as they are pronounced, plus tone if it is likely to be pronounced in different tones (use standard pinyin 4-tone annotation). The reason is that some fillers, when pronounced in a different tone, are used as back-channels and discourse markers. Common mandarin pause fillers are: eh1, en1, em1 and neige (if it is pronounced as "nage" then write "nage"). Use your intuition to judge whether "neige" is used as a pause filler or a real demonstrative.

Try to be accurate with whether there is a nasal consonant at the end of the pause filler, e.g. whether it is an "eh" or "en". The vowel choice is less important. Note also that pause fillers are often lengthened and preceded/followed by pauses. Don't forget to use standard transcription style in section 5 to transcribe lengthening and pauses.

For other editing terms, such as {就是} {或者是}, simply use curly brackets to surround them.

11.2 Disfluencies

Please follow the general instructions in Section 5 on disfluency annotation. Here are a few examples.

- Here is an example of nested repair:
 - 就感觉客厅是((公+公+公)+公共:)就是休息啊或者接待客人的地方
- Here is an example with several editing terms:
 - 我觉得书房的话可以弄得就是:是那种背对背式((或者是:+对就是:或者是)(相互+就是:hm面对面式的那种))
- Here is an example with abandoned partial utterance at the end:
 - 客厅:可能我觉得要好好不知一下吧(因为+)

11.3 Pronunciation variations and words in languages other than Chinese

Ignore variations in pronunciation due to regional influences. In the case of "儿话音" where an "er" added to the end of a word, write "儿" in the Chinese transcription, e.g. "小碎花儿". For words other than Chinese, if it is conventionalized, use the conventional orthography. E.g. the use of "ing" after Chinese verb, such as "吃饭ing", just write "ing" after the Chinese word. Otherwise use the orthography of the target language.

11.4 English translation tiers

For English translation, write a paraphrase to the best of our ability in English in the appropriate ch_en tier. We ask you to use standard English spelling and to preserve the disfluency markings you have made in Chinese, using the same bracketing scheme as in Section 5. For pause fillers, maintain the original orthography, e.g. F en1 F neige. Make sure that they are in Pinyin and not in Chinese characters.

References

- Paul Boersma. 2002. Praat, a system for doing phonetics by computer. *Glott international*, 5(9/10):341–345.
- M. Meteer, A. Taylor, R. MacIntyre, and R. Iyer. 1995. Disfluency annotation stylebook for the switchboard corpus. ms. Technical report, Department of Computer and Information Science, University of Pennsylvania. URL <ftp://ftp.cis.upenn.edu/pub/treebank/swbd/doc/DFL-book.ps>.
- David Schlangen and Raquel Fernández. 2008. The potsdam dialogue corpora: Transcription and annotation manual.